

NVM Express Technical Errata

Errata ID	020
Change Date	8/17/2011
Affected Spec Ver.	NVM Express 1.0b
Corrected Spec Ver.	

Submission info

Name	Company	Date
Peter Onufryk	IDT	8/17/2011
Kevin Marks	Dell	8/17/2011
Ken Okin	Virident	8/17/2011

This erratum adds a definition of LBA range.

The behavior when a bad doorbell value is written is clarified.

Editorial changes are made to sections 1, 5, and 6.

Description of the specification technical flaw

Add section 1.6.12 as shown below:

1.6.12 LBA range

A collection of contiguous logical blocks specified by a starting LBA and a number of logical blocks.

Modify the second to last paragraph in section 4.1 as shown below:

If host software writes an invalid value to the Submission Queue Tail Doorbell or Completion Queue Head Doorbell register and an Asynchronous Event Request command is outstanding, then an asynchronous event is posted to the Admin Completion Queue with a status code of Invalid Doorbell Write Value. **The associated queue should be deleted and recreated by host software. For a Submission Queue that experiences this error, the controller may complete previously fetched commands; no additional commands are fetched.** This condition may be caused by host software attempting to add an entry to a full Submission Queue or remove an entry from an empty Completion Queue. **(ADD RETURN)**

The behavior if a command is overwritten is undefined.

Update Figure 57 as shown below:

Figure 57: Get Log Page – Log Page Identifiers

Log Identifier	O/M	Description
00h		Reserved
01h	M	Error Information
02h	M	SMART / Health Information
03h	M	Firmware Slot Information
04h – 7Fh		Reserved
80h – BFh		I/O Command Set Specific
C0h – FFh		Vendor specific

O/M: O = Optional, M = Mandatory

Update the last two paragraphs of section 5.10.1.1 as shown below:

This error log may return the last n errors. If **the** host software specifies a data transfer of the size of n error logs, then the error logs for the last n errors is returned. The ordering of the entries is based on the time when the error occurred, with the most recent error being returned as the first log.

Each entry in the log page returned is defined in Figure 58. The log page is a set of 64 byte entries; the number of entries supported is indicated in the Identify Controller data structure **in Figure 65**.

Modify the first two paragraphs of section 5.10.1.2 as shown below:

This log page is used to provide SMART and general health information. The information provided is over the life of the controller and is retained across power cycles. The log page shall be supported on a global basis. The log page may be supported on a per namespace basis, as indicated in the Identify Controller data structure in Figure 65. To request the global log page, the namespace specified is FFFFFFFFh.

Critical warnings regarding the health of the device may be indicated via an asynchronous event notification to the host. The warnings that result in an asynchronous event notification to the host are ~~may-be~~ configured using the Set Features command; refer to section 5.12.1.11.

Modify section 5.10.2 as shown below:

A completion queue entry is posted to the Admin Completion Queue ~~when if~~ the log has been transferred posted to the memory buffer indicated in PRP Entry 1. Get Log Page command specific errors are defined in Figure 61.

Modify the first paragraph of section 5.11 as shown below:

The Identify command returns a data buffer that describes the controller or namespace capabilities and status. The data structure is 4096 bytes in size. The host indicates specifies as a command parameter whether to return the controller or namespace specific data structure. For the namespace data structure, the definition of the structure is specific to the I/O Command Set command set selected for use.

The Identify command uses the PRP Entry 1, PRP Entry 2, and Command Dword 10 fields. All other command specific fields are reserved.

Modify section 5.11.1 as shown below:

A completion queue entry is posted to the Admin Completion Queue ~~when if~~ the Identify data structure has been transferred posted to the memory buffer indicated in PRP Entry 1.

Make corrections to Table of Contents:

- Add section 1.6 “Definitions”.
- Adjust indentation level of section 2.7 “Other Capability Registers”.
- Adjust indentation level of section 9.4 “Internal Device Error Handling” and 9.5 “Controller Fatal Status Condition”

In section 2.3.4, change font size to 9 point in table.

Add empty line prior to section 1.9.

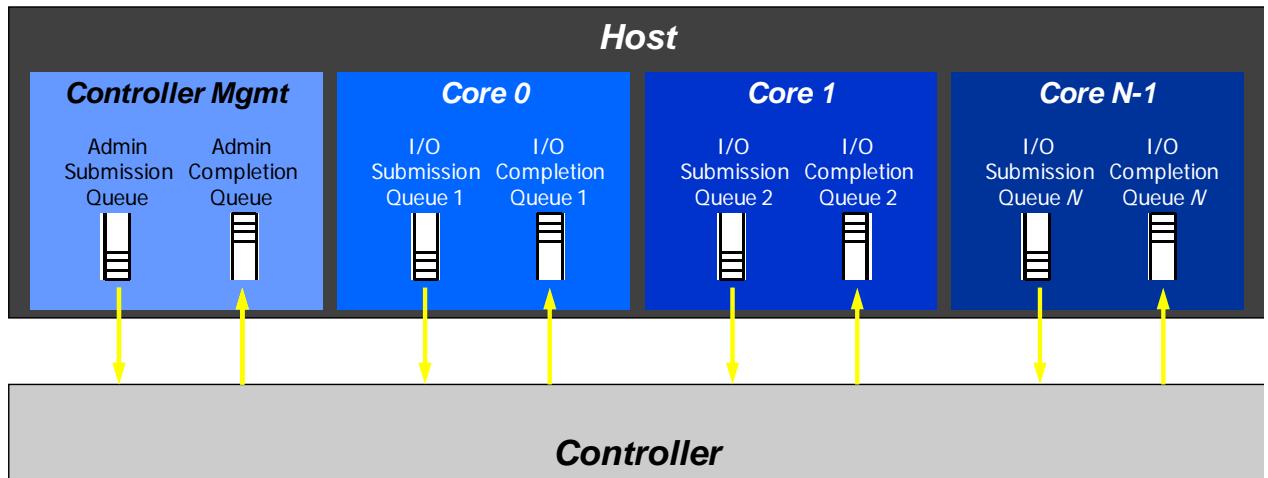
Modify the title of Figure 3 as shown:

Figure 3: Byte, word and Dword ~~relationships~~ Relationships

Modify the table in section 2.1 as shown:

1Ch	1Fh 20h	BAR3	Reserved <BAR3>
20h	23h	BAR4	Vendor Specific

Update Figure 1 as shown (changing “Core N” to “Core N-1”):



Update section 6 as shown below:

The device is comprised of some number of namespaces, where each namespace is comprised of some number of logical blocks. A logical block is the smallest unit of data that may be read or written from the controller. The logical block data size, reported in bytes, is always a power of two. **LBA Logical block** sizes may be 512 bytes, 1KB, 2KB, 4KB, 8KB, etc. Supported **LBA logical block** sizes are reported in the Identify Namespace data structure.

The NVM **Command Set command-set** includes the commands listed in Figure 98. The following subsections describe the definition for each of these commands. Commands shall only be issued by the host when the controller is ready as indicated in the Controller Status register (CSTS.RDY) and after appropriate I/O Submission Queue(s) and I/O Completion Queue(s) have been created.

The Submission Queue Entry (SQE) structure and the fields that are common to all NVM commands are defined in section 4.2. The Completion Queue Entry (CQE) structure and the fields that are common to all NVM commands are defined in section 4.5. The command specific fields in the SQE and CQE structures for the NVM Command Set are defined in this section.

In the case of Compare, Read, and Write commands, the host may indicate whether a time limit should be applied to the operation by setting the Limited Retry (LR) field in the command. The time limit is **specified indicated** in the Error Recovery feature, specified in section 5.12.1.5. If the host does not **specify indicate** a time limit should be applied, then the controller should apply all error recovery means to complete the operation.

Modify Figure 98 as shown below:

Figure 98: Opcodes for NVM Commands

Opcode (07)	Opcode (06:02)	Opcode (01:00)	Opcode ²	O/M ¹	Command ³
Standard Command	Function	Data Transfer			
0b	000 00b	00b	00h	M	Flush
0b	000 00b	01b	01h	M	Write
0b	000 00b	10b	02h	M	Read
0b	000 01b	00b	04h	O	Write Uncorrectable
0b	000 01b	01b	05h	O	Compare
0b	000 10b	01b	09h	O	Dataset Management
Vendor Specific					
1b	na	na	80h – FFh	O	Vendor specific

NOTES:

1. O/M definition: O = Optional, M = Mandatory.
2. Opcodes not listed are reserved.
3. All NVM commands use the Namespace Identifier field (CDW1.NSID).

Modify the third paragraph of section 6.1 as shown below:

A namespace may or may not have a relationship to a Submission Queue; this relationship is determined by the **host** software implementation. The controller shall support access to any valid namespace from any I/O Submission Queue.

Modify the first paragraph of section 6.2 as shown below:

The command sequences that may be used in a fused operation for the NVM Command Set are defined in Figure 99, ~~the host shall not construct other fused operation sequences for the NVM Command Set~~. Refer to section 4.6 for details on fused operations.

Modify the section 6.2.1 as shown below:

The Compare and Write fused operation compares the contents of the **LBA(s) logical block(s) specified indicated** in the Compare command to the data stored at the **indicated specified** LBA range. If the compare is successful, then the LBA range is updated with the data provided in the Write command. If the Compare operation is not successful, then the Write operation is aborted with a status of Command Aborted due to Failed Fused Command and the **contents in the LBA range are is** not modified. If the Write operation is not successful, the Compare operation completion status is unaffected.

Disposition log

8/17/2011	Erratum partially captured.
8/31/2011	Additional material included.
10/3/2011	Erratum ratified.

Technical input submitted to the NVMHCI Workgroup is subject to the terms of the NVMHCI Contributor's agreement.